# Bayes Coffee House: Probabilistic Bisimulation Metrics and Their Applications to Representation Learning

## Lecture 1: The logical characterization of bisimulation

Prakash Panangaden
School of Computer Science, McGill University
Montreal Institute of Learning Algorithms
School of Informatics, The University of Edinburgh

22nd February 2023, Edinburgh

# Overview

- Discrete probabilistic transition system.

# Overview

- Discrete probabilistic transition system.
- Labelled Markov processes: probabilistic transition systems with continuous state spaces.

# Overview

- Discrete probabilistic transition system.
- Labelled Markov processes: probabilistic transition systems with continuous state spaces.
- Bisimulation for LMPs.

# Overview

- Discrete probabilistic transition system.
- Labelled Markov processes: probabilistic transition systems with continuous state spaces.
- Bisimulation for LMPs.
- A game for bisimulation.

# Overview

- Discrete probabilistic transition system.
- Labelled Markov processes: probabilistic transition systems with continuous state spaces.
- Bisimulation for LMPs.
- A game for bisimulation.
- Simulation

# Overview

- Discrete probabilistic transition system.
- Labelled Markov processes: probabilistic transition systems with continuous state spaces.
- Bisimulation for LMPs.
- A game for bisimulation.
- Simulation
- Logical characterization.

# Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS97]

# Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS97]
- Logical characterization. [LICS98,Info and Comp 2002]

# Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS97]
- Logical characterization. [LICS98,Info and Comp 2002]
- Metric analogue of bisimulation. [CONCUR99, TCS2004]

# Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS97]
- Logical characterization. [LICS98,Info and Comp 2002]
- Metric analogue of bisimulation. [CONCUR99, TCS2004]
- Approximation of LMPs. [LICS00,Info and Comp 2003]

# Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS97]
- Logical characterization. [LICS98,Info and Comp 2002]
- Metric analogue of bisimulation. [CONCUR99, TCS2004]
- Approximation of LMPs. [LICS00,Info and Comp 2003]
- Weak bisimulation. [LICS02,CONCUR02]

# Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS97]
- Logical characterization. [LICS98,Info and Comp 2002]
- Metric analogue of bisimulation. [CONCUR99, TCS2004]
- Approximation of LMPs. [LICS00,Info and Comp 2003]
- Weak bisimulation. [LICS02,CONCUR02]
- Real time. [QEST 2004, JLAP 2003,LMCS 2006]

# Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS97]
- Logical characterization. [LICS98,Info and Comp 2002]
- Metric analogue of bisimulation. [CONCUR99, TCS2004]
- Approximation of LMPs. [LICS00,Info and Comp 2003]
- Weak bisimulation. [LICS02,CONCUR02]
- Real time. [QEST 2004, JLAP 2003,LMCS 2006]
- Event bisimulation [Info and Comp 2006]

# Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS97]
- Logical characterization. [LICS98,Info and Comp 2002]
- Metric analogue of bisimulation. [CONCUR99, TCS2004]
- Approximation of LMPs. [LICS00,Info and Comp 2003]
- Weak bisimulation. [LICS02,CONCUR02]
- Real time. [QEST 2004, JLAP 2003,LMCS 2006]
- Event bisimulation [Info and Comp 2006]
- Metrics for MDPs [UAI 2004,2005, SIAM 2011]

# Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS97]
- Logical characterization. [LICS98,Info and Comp 2002]
- Metric analogue of bisimulation. [CONCUR99, TCS2004]
- Approximation of LMPs. [LICS00,Info and Comp 2003]
- Weak bisimulation. [LICS02,CONCUR02]
- Real time. [QEST 2004, JLAP 2003,LMCS 2006]
- Event bisimulation [Info and Comp 2006]
- Metrics for MDPs [UAI 2004,2005, SIAM 2011]
- Approximation by averaging [ICALP 2009, JACM 2014]

# Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS97]
- Logical characterization. [LICS98,Info and Comp 2002]
- Metric analogue of bisimulation. [CONCUR99, TCS2004]
- Approximation of LMPs. [LICS00,Info and Comp 2003]
- Weak bisimulation. [LICS02,CONCUR02]
- Real time. [QEST 2004, JLAP 2003,LMCS 2006]
- Event bisimulation [Info and Comp 2006]
- Metrics for MDPs [UAI 2004,2005, SIAM 2011]
- Approximation by averaging [ICALP 2009, JACM 2014]
- Duality [JACM 2014, LICS 2013, 2017]

# Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS97]
- Logical characterization. [LICS98,Info and Comp 2002]
- Metric analogue of bisimulation. [CONCUR99, TCS2004]
- Approximation of LMPs. [LICS00,Info and Comp 2003]
- Weak bisimulation. [LICS02,CONCUR02]
- Real time. [QEST 2004, JLAP 2003,LMCS 2006]
- Event bisimulation [Info and Comp 2006]
- Metrics for MDPs [UAI 2004,2005, SIAM 2011]
- Approximation by averaging [ICALP 2009, JACM 2014]
- Duality [JACM 2014, LICS 2013, 2017]
- Quantitative equational logic [LICS 2016, 2017, 2018, 2021, CALCO 2021]

# Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS97]
- Logical characterization. [LICS98,Info and Comp 2002]
- Metric analogue of bisimulation. [CONCUR99, TCS2004]
- Approximation of LMPs. [LICS00,Info and Comp 2003]
- Weak bisimulation. [LICS02,CONCUR02]
- Real time. [QEST 2004, JLAP 2003,LMCS 2006]
- Event bisimulation [Info and Comp 2006]
- Metrics for MDPs [UAI 2004,2005, SIAM 2011]
- Approximation by averaging [ICALP 2009, JACM 2014]
- Duality [JACM 2014, LICS 2013, 2017]
- Quantitative equational logic [LICS 2016, 2017, 2018, 2021, CALCO 2021]
- Diffusion and continuous-time processes [MFPS 2019, 2020]

# Collaborators

Giorgio Bacci, Philippe Chaput, Linan Chen, Florence Clerc, Vincent Danos, Josée Desharnais, Abbas Edalat, Norm Ferns, Nathanaël Fijalkow, Robert Furber, Vineet Gupta, Radha Jagadeesan, Bartek Klin, Dexter Kozen, Kim Larsen, François Laviolette, Radu Mardare, Gordon Plotkin and Doina Precup.

# Labelled Transition System

- A set of states $S$,

# Labelled Transition System

- A set of states $S$,
- a set of *labels* or *actions*, L or $\mathcal{A}$ and

# Labelled Transition System

- A set of states $S$,
- a set of *labels* or *actions*, L or $\mathcal{A}$ and
- a transition relation $\subseteq S \times \mathcal{A} \times S$, usually written

$$\rightarrow_a \subseteq S \times S.$$

The transitions could be indeterminate (nondeterministic).

# Markov Chains

- A *discrete-time* Markov chain is a finite set $S$ (the state space) together with a transition probability function $T : S \times S \longrightarrow [0, 1]$.

# Markov Chains

- A *discrete-time* Markov chain is a finite set $S$ (the state space) together with a transition probability function $T : S \times S \longrightarrow [0, 1]$.
- The key property is that the transition probability from $s$ to $s'$ only depends on $s$ and $s'$ and not on the past history of how it got there. This is what allows the probabilistic data to be given as a single matrix $T$.

# Discrete probabilistic transition systems

- Just like a labelled transition system with probabilities associated with the transitions.

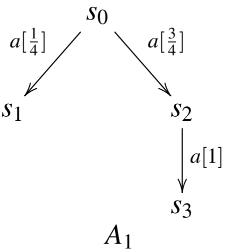
# Discrete probabilistic transition systems

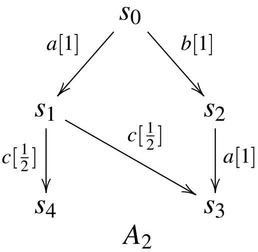
- Just like a labelled transition system with probabilities associated with the transitions.
- $$(S, \mathsf{L}, \forall a \in \mathsf{L}\ T_a : S \times S \longrightarrow [0, 1])$$

# Discrete probabilistic transition systems

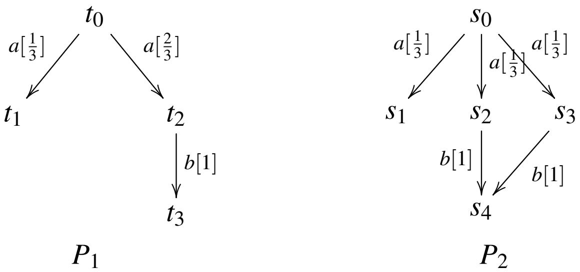
- Just like a labelled transition system with probabilities associated with the transitions.
- $$(S, \mathsf{L}, \forall a \in \mathsf{L}\ T_a : S \times S \longrightarrow [0,1])$$
- The model is *reactive*: All probabilistic data is *internal* - no probabilities associated with environment behaviour.

# Examples of PTSs

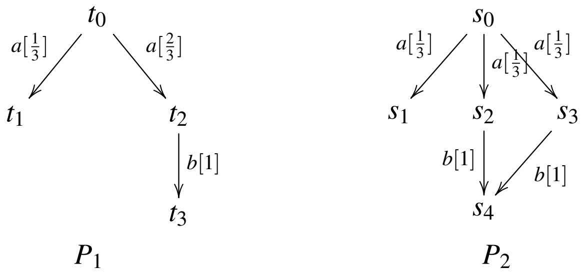


- Consider


$t_0$
$a[\frac{1}{3}]$
$a[\frac{2}{3}]$
$t_1$
$t_2$
$b[1]$
$t_3$
$P_1$
$s_0$
$a[\frac{1}{3}]$
$a[\frac{1}{3}]$
$a[\frac{1}{3}]$
$s_1$
$s_2$
$s_3$
$b[1]$
$b[1]$
$s_4$
$P_2$

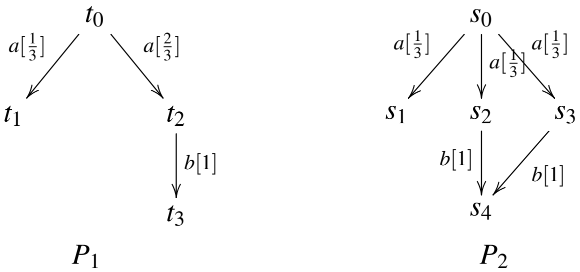
- Consider




- Should $s_0$ and $t_0$ be bisimilar?

# Bisimulation for PTS: Larsen and Skou

- Consider




- Should $s_0$ and $t_0$ be bisimilar?
- Yes, but we need to add the probabilities.

# The Official Definition

- Let $\mathcal{S} = (S, \mathsf{L}, T_a)$ be a PTS. An equivalence relation $R$ on $S$ is a **bisimulation** if whenever $sRs'$, with $s, s' \in S$, we have that for all $a \in \mathcal{A}$ and every $R$-equivalence class, $A$, $T_a(s, A) = T_a(s', A)$.

# The Official Definition

- Let $\mathcal{S} = (S, \mathsf{L}, T_a)$ be a PTS. An equivalence relation $R$ on $S$ is a **bisimulation** if whenever $sRs'$, with $s, s' \in S$, we have that for all $a \in \mathcal{A}$ and every $R$-equivalence class, $A$, $T_a(s, A) = T_a(s', A)$.
- The notation $T_a(s, A)$ means “the probability of starting from $s$ and jumping to a state in the set $A$.”

# The Official Definition

- Let $\mathcal{S} = (S, \mathsf{L}, T_a)$ be a PTS. An equivalence relation $R$ on $S$ is a **bisimulation** if whenever $sRs'$, with $s, s' \in S$, we have that for all $a \in \mathcal{A}$ and every $R$-equivalence class, $A$, $T_a(s, A) = T_a(s', A)$.
- The notation $T_a(s, A)$ means "the probability of starting from $s$ and jumping to a state in the set $A$."
- Two states are bisimilar if there is some bisimulation relation $R$ relating them.

# What are labelled Markov processes?

- Labelled Markov processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is governed by a probability distribution - no other indeterminacy.

# What are labelled Markov processes?

- Labelled Markov processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is governed by a probability distribution - no other indeterminacy.
- All probabilistic data is *internal* - no probabilities associated with environment behaviour.

# What are labelled Markov processes?

- Labelled Markov processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is governed by a probability distribution - no other indeterminacy.
- All probabilistic data is *internal* - no probabilities associated with environment behaviour.
- We observe the interactions - not the internal states.

# What are labelled Markov processes?

- Labelled Markov processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is governed by a probability distribution - no other indeterminacy.
- All probabilistic data is *internal* - no probabilities associated with environment behaviour.
- We observe the interactions - not the internal states.
- **In general, the state space of a labelled Markov process may be a *continuum*.**

# The Need for Measure Theory

- Basic fact: There are subsets of $\mathbf{R}$ for which no sensible notion of size can be defined.

# The Need for Measure Theory

- Basic fact: There are subsets of $\mathbf{R}$ for which no sensible notion of size can be defined.
- More precisely, there is no non-trivial translation-invariant measure defined on all the subsets of the reals.

# Stochastic Kernels

- A *stochastic kernel* (Markov kernel) is a function $h : S \times \Sigma \longrightarrow [0, 1]$ with (a) $h(s, \cdot) : \Sigma \longrightarrow [0, 1]$ a (sub)probability measure and (b) $h(\cdot, A) : S \longrightarrow [0, 1]$ a measurable function.

# Stochastic Kernels

- A *stochastic kernel* (Markov kernel) is a function $h : S \times \Sigma \longrightarrow [0, 1]$ with (a) $h(s, \cdot) : \Sigma \longrightarrow [0, 1]$ a (sub)probability measure and (b) $h(\cdot, A) : S \longrightarrow [0, 1]$ a measurable function.
- Though apparently asymmetric, these are the stochastic analogues of binary relations

# Stochastic Kernels

- A *stochastic kernel* (Markov kernel) is a function $h : S \times \Sigma \longrightarrow [0, 1]$ with (a) $h(s, \cdot) : \Sigma \longrightarrow [0, 1]$ a (sub)probability measure and (b) $h(\cdot, A) : S \longrightarrow [0, 1]$ a measurable function.
- Though apparently asymmetric, these are the stochastic analogues of binary relations
- and the uncountable generalization of a matrix.

# Stochastic Kernels

- A *stochastic kernel* (Markov kernel) is a function $h : S \times \Sigma \longrightarrow [0, 1]$ with (a) $h(s, \cdot) : \Sigma \longrightarrow [0, 1]$ a (sub)probability measure and (b) $h(\cdot, A) : S \longrightarrow [0, 1]$ a measurable function.
- Though apparantly asymmetric, these are the stochastic analogues of binary relations
- and the uncountable generalization of a matrix.
- They are the Kleisli arrows of a monad: the Giry monad.

# Formal Definition of LMPs

- An LMP is a tuple $(S, \Sigma, \mathsf{L}, \forall\alpha \in \mathsf{L}.\tau_\alpha)$ where $\tau_\alpha : S \times \Sigma \longrightarrow [0, 1]$ is a *transition probability* function such that

# Formal Definition of LMPs

- An LMP is a tuple $(S, \Sigma, \mathsf{L}, \forall\alpha \in \mathsf{L}.\tau_\alpha)$ where $\tau_\alpha : S \times \Sigma \longrightarrow [0, 1]$ is a *transition probability* function such that
- $\forall s : S.\lambda A : \Sigma.\tau_\alpha(s, A)$ is a subprobability measure
  and
  $\forall A : \Sigma.\lambda s : S.\tau_\alpha(s, A)$ is a measurable function.

# Probabilistic Bisimulation

## Desharnais et al.

Let $\mathcal{S} = (S, i, \Sigma, \tau)$ be a labelled Markov process. An equivalence relation $R$ on $S$ is a **bisimulation** if whenever $sRs'$, with $s, s' \in S$, we have that for all $a \in \mathcal{A}$ and every $R$-closed measurable set $A \in \Sigma$, $\tau_a(s, A) = \tau_a(s', A)$.

Two states are bisimilar if they are related by a bisimulation relation.

# A game for bisimulation

- Two players: spoiler (S) and duplicator (D).

# A game for bisimulation

- Two players: spoiler (S) and duplicator (D).
- Duplicator claims $x, y$ are bisimilar.

# A game for bisimulation

- Two players: spoiler (S) and duplicator (D).
- Duplicator claims $x, y$ are bisimilar.
- Spoiler exhibits a set $C$ and says $C$ is bisimulation-closed and that $\tau(x, C) \neq \tau(y, C)$. Assume that the inequality holds; it is easy to check.

# A game for bisimulation

- Two players: spoiler (S) and duplicator (D).
- Duplicator claims $x, y$ are bisimilar.
- Spoiler exhibits a set $C$ and says $C$ is bisimulation-closed and that $\tau(x, C) \neq \tau(y, C)$. Assume that the inequality holds; it is easy to check.
- Duplicator responds by saying that $C$ is not bisimulation-closed and that exhibits $x' \in C$ and $y' \notin C$ and claims that $x', y'$ are bisimilar.

# A game for bisimulation

- Two players: spoiler (S) and duplicator (D).
- Duplicator claims $x, y$ are bisimilar.
- Spoiler exhibits a set $C$ and says $C$ is bisimulation-closed and that $\tau(x, C) \neq \tau(y, C)$. Assume that the inequality holds; it is easy to check.
- Duplicator responds by saying that $C$ is not bisimulation-closed and that exhibits $x' \in C$ and $y' \notin C$ and claims that $x', y'$ are bisimilar.
- A player loses when he or she cannot make a move. Note that if $C$ is all of the state space, duplicator loses. Duplicator wins if she can play forever.

# A game for bisimulation

- Two players: spoiler (S) and duplicator (D).
- Duplicator claims $x, y$ are bisimilar.
- Spoiler exhibits a set $C$ and says $C$ is bisimulation-closed and that $\tau(x, C) \neq \tau(y, C)$. Assume that the inequality holds; it is easy to check.
- Duplicator responds by saying that $C$ is not bisimulation-closed and that exhibits $x' \in C$ and $y' \notin C$ and claims that $x', y'$ are bisimilar.
- A player loses when he or she cannot make a move. Note that if $C$ is all of the state space, duplicator loses. Duplicator wins if she can play forever.
- We prove that $x$ is bisimilar to $y$ iff Duplicator has a winning strategy starting from $(x, y)$.

# Logical Characterization

- $$\mathcal{L} ::== \mathsf{T} | \phi_1 \wedge \phi_2 | \langle a \rangle_q \phi$$

# Logical Characterization

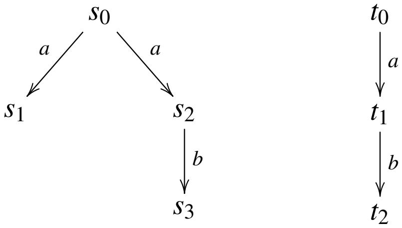
- $$\mathcal{L} ::== \mathsf{T} | \phi_1 \wedge \phi_2 | \langle a \rangle_q \phi$$
- We say $s \models \langle a \rangle_q \phi$ iff

$$\exists A \in \Sigma.(\forall s' \in A.s' \models \phi) \wedge (\tau_a(s, A) > q).$$

# Logical Characterization

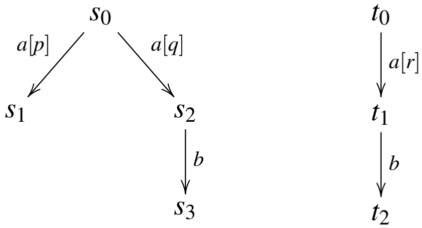
- $$\mathcal{L} ::== \mathsf{T} | \phi_1 \wedge \phi_2 | \langle a \rangle_q \phi$$
- We say $s \models \langle a \rangle_q \phi$ iff
$$\exists A \in \Sigma.(\forall s' \in A. s' \models \phi) \wedge (\tau_a(s, A) > q).$$
- Two systems are bisimilar iff they obey the same formulas of $\mathcal{L}$. [DEP 1998 LICS, I and C 2002]

# That cannot be right?



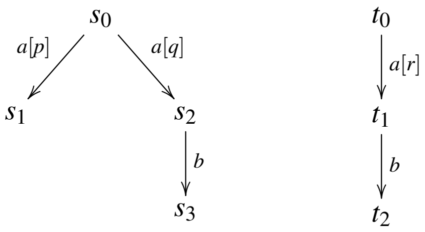

Two processes that cannot be distinguished without negation.
The formula that distinguishes them is $\langle a \rangle (\neg \langle b \rangle \top)$.

# But it is!

- 

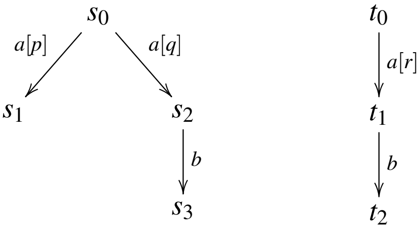

We add probabilities to the transitions.

# But it is!

- 


  We add probabilities to the transitions.
- If $p + q < r$ or $p + q > r$ we can easily distinguish them.

# But it is!

- 


  We add probabilities to the transitions.
- If $p + q < r$ or $p + q > r$ we can easily distinguish them.
- If $p + q = r$ and $p > 0$ then $q < r$ so $\langle a \rangle r \langle b \rangle 1 \top$ distinguishes them.

# Proof idea

- Show that the relation "$s$ and $s'$ satisfy exactly the same formulas" is a bisimulation.

# Proof idea

- Show that the relation "$s$ and $s'$ satisfy exactly the same formulas" is a bisimulation.
- Can easily show that $\tau_a(s, A) = \tau_a(s', A)$ for $A$ of the form $[\![\phi]\!]$.

# Proof idea

- Show that the relation "$s$ and $s'$ satisfy exactly the same formulas" is a bisimulation.
- Can easily show that $\tau_a(s, A) = \tau_a(s', A)$ for $A$ of the form $[\![\phi]\!]$.
- Use Dynkin's lemma to show that we get a well defined measure on the $\sigma$-algebra generated by such sets and the above equality holds.

# Proof idea

- Show that the relation "$s$ and $s'$ satisfy exactly the same formulas" is a bisimulation.
- Can easily show that $\tau_a(s, A) = \tau_a(s', A)$ for $A$ of the form $[\![\phi]\!]$.
- Use Dynkin's lemma to show that we get a well defined measure on the $\sigma$-algebra generated by such sets and the above equality holds.
- Use special properties of analytic spaces to show that this $\sigma$-algebra is the same as the original $\sigma$-algebra.

# Simulation

Let $\mathcal{S} = (S, \Sigma, \tau)$ be a labelled Markov process. A preorder $R$ on $S$ is a **simulation** if whenever $sRs'$, we have that for all $a \in \mathcal{A}$ and every $R$-closed measurable set $A \in \Sigma$, $\tau_a(s, A) \leq \tau_a(s', A)$. We say $s$ is simulated by $s'$ if $sRs'$ for some simulation relation $R$.

# Logic for simulation?

- The logic used in the characterization has no negation, not even a limited negative construct.

# Logic for simulation?

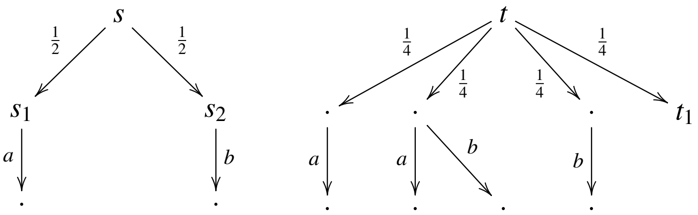
- The logic used in the characterization has no negation, not even a limited negative construct.
- One can show that if $s$ simulates $s'$ then $s$ satisfies all the formulas of $\mathcal{L}$ that $s'$ satisfies.

# Logic for simulation?

- The logic used in the characterization has no negation, not even a limited negative construct.
- One can show that if $s$ simulates $s'$ then $s$ satisfies all the formulas of $\mathcal{L}$ that $s'$ satisfies.
- What about the converse?

# Counter example!

In the following picture, $t$ satisfies all formulas of $\mathcal{L}$ that $s$ satisfies but $t$ does not simulate $s$.




All transitions from $s$ and $t$ are labelled by $a$.

# Counter example (contd.)

- A formula of $\mathcal{L}$ that is satisfied by $t$ but not by $s$.

$$\langle a\rangle_0(\langle a\rangle_0 \mathsf{T} \wedge \langle b\rangle_0 \mathsf{T}).$$

# Counter example (contd.)

- A formula of $\mathcal{L}$ that is satisfied by $t$ but not by $s$.

$$\langle a\rangle_0(\langle a\rangle_0 \mathsf{T} \wedge \langle b\rangle_0 \mathsf{T}).$$

- A formula with disjunction that is satisfied by $s$ but not by $t$:

$$\langle a\rangle_{\frac{3}{4}}(\langle a\rangle_0 \mathsf{T} \vee \langle b\rangle_0 \mathsf{T}).$$

# A logical characterization for simulation

- The logic $\mathcal{L}$ does **not** characterize simulation. One needs disjunction.

$$\mathcal{L}_\vee := \mathcal{L} \mid \phi_1 \vee \phi_2.$$

# A logical characterization for simulation

- The logic $\mathcal{L}$ does **not** characterize simulation. One needs disjunction.

$$\mathcal{L}_\vee := \mathcal{L} \mid \phi_1 \vee \phi_2.$$

- With this logic we have:
  An **LMP** $s_1$ simulates $s_2$ if and only if for every formula $\phi$ of $\mathcal{L}_\vee$ we have

$$s_1 \models \phi \Rightarrow s_2 \models \phi.$$

# A logical characterization for simulation

- The logic $\mathcal{L}$ does **not** characterize simulation. One needs disjunction.

$$\mathcal{L}_\vee := \mathcal{L} \mid \phi_1 \vee \phi_2.$$

- With this logic we have:
  An **LMP** $s_1$ simulates $s_2$ if and only if for every formula $\phi$ of $\mathcal{L}_\vee$ we have

$$s_1 \models \phi \Rightarrow s_2 \models \phi.$$

- The original proof uses domain theory and approximation.

# A logical characterization for simulation

- The logic $\mathcal{L}$ does **not** characterize simulation. One needs disjunction.

$$\mathcal{L}_\vee := \mathcal{L} \mid \phi_1 \vee \phi_2.$$

- With this logic we have:
  An **LMP** $s_1$ simulates $s_2$ if and only if for every formula $\phi$ of $\mathcal{L}_\vee$ we have

$$s_1 \models \phi \Rightarrow s_2 \models \phi.$$

- The original proof uses domain theory and approximation.
- New development (2017 ICALP) we can prove logical characterization for simulation and bisimulation in almost the same way.

# Digression on Analytic Spaces

- An analytic set $A$ is the image of a Polish space $X$ (or a Borel subset of $X$) under a continuous (or measurable) function $f : X \to Y$, where $Y$ is Polish. If $(S, \Sigma)$ is a measurable space where $S$ is an analytic set in some ambient topological space and $\Sigma$ is the Borel $\sigma$-algebra on $S$.

# Digression on Analytic Spaces

- An analytic set $A$ is the image of a Polish space $X$ (or a Borel subset of $X$) under a continuous (or measurable) function $f : X \to Y$, where $Y$ is Polish. If $(S, \Sigma)$ is a measurable space where $S$ is an analytic set in some ambient topological space and $\Sigma$ is the Borel $\sigma$-algebra on $S$.
- Analytic sets do not form a $\sigma$-algebra but they are in the completion of the Borel algebra under **any** measure. [Universally measurable.]

# Amazing Facts about Analytic Spaces

- Given $A$ an analytic space and $\sim$ an equivalence relation such that there is a *countable* family of real-valued measurable functions $f_i : S \longrightarrow \mathbf{R}$ such that

$$\forall s, s' \in S. s \sim s' \iff \forall f_i. f_i(s) = f_i(s')$$

then the quotient space $(Q, \Omega)$ - where $Q = S/\sim$ and $\Omega$ is the finest $\sigma$-algebra making the canonical surjection $q : S \longrightarrow Q$ measurable - is also analytic.

# Amazing Facts about Analytic Spaces

- Given $A$ an analytic space and $\sim$ an equivalence relation such that there is a *countable* family of real-valued measurable functions $f_i : S \longrightarrow \mathbf{R}$ such that

$$\forall s, s' \in S. s \sim s' \iff \forall f_i . f_i(s) = f_i(s')$$

  then the quotient space $(Q, \Omega)$ - where $Q = S/\sim$ and $\Omega$ is the finest $\sigma$-algebra making the canonical surjection $q : S \longrightarrow Q$ measurable - is also analytic.
- If an analytic space $(S, \Sigma)$ has a sub-$\sigma$-algebra $\Sigma_0$ of $\Sigma$ which separates points and is countably generated then $\Sigma_0$ is $\Sigma$! The Unique Structure Theorem (UST).

# Some more measure theory

- A $\pi$-system is a family of sets closed under finite intersections.

# Some more measure theory

- A $\pi$-system is a family of sets closed under finite intersections.
- A $\lambda$-system is a family of sets closed under complements and countable *disjoint* unions.

# Some more measure theory

- A $\pi$-system is a family of sets closed under finite intersections.
- A $\lambda$-system is a family of sets closed under complements and countable *disjoint* unions.
- $\lambda - \pi$ theorem: If $\Pi$ is a $\pi$-system and $\Lambda$ is a $\lambda$-system and $\Pi \subset \Lambda$ then $\sigma(\Pi) \subset \Lambda$.

# Some more measure theory

- A $\pi$-system is a family of sets closed under finite intersections.
- A $\lambda$-system is a family of sets closed under complements and countable *disjoint* unions.
- $\lambda - \pi$ theorem: If $\Pi$ is a $\pi$-system and $\Lambda$ is a $\lambda$-system and $\Pi \subset \Lambda$ then $\sigma(\Pi) \subset \Lambda$.
- Corollary: If two measures agree on the sets of a $\pi$-system then they agree on the generated $\sigma$-algebra.

# Bisimulation proof I

- Given $(S, \Sigma, \tau_a)$ an LMP, we define $x \simeq y$ if $x$ and $y$ obey exactly the same formulas of $\mathcal{L}_0$.

# Bisimulation proof I

- Given $(S, \Sigma, \tau_a)$ an LMP, we define $x \simeq y$ if $x$ and $y$ obey exactly the same formulas of $\mathcal{L}_0$.
- We claim that $\simeq$ is a bisimulation relation.

# Bisimulation proof I

- Given $(S, \Sigma, \tau_a)$ an LMP, we define $x \simeq y$ if $x$ and $y$ obey exactly the same formulas of $\mathcal{L}_0$.
- We claim that $\simeq$ is a bisimulation relation.
- Suppose that $x, y \in S$ and for some $a$ and some $\simeq$-closed set $C$, $\tau_a(x, C) \neq \tau_a(y, C)$.

# Bisimulation proof I

- Given $(S, \Sigma, \tau_a)$ an LMP, we define $x \simeq y$ if $x$ and $y$ obey exactly the same formulas of $\mathcal{L}_0$.
- We claim that $\simeq$ is a bisimulation relation.
- Suppose that $x, y \in S$ and for some $a$ and some $\simeq$-closed set $C$, $\tau_a(x, C) \neq \tau_a(y, C)$.
- We need to show there is a formula on which $x, y$ disagree.

# Bisimulation proof I

- Given $(S, \Sigma, \tau_a)$ an LMP, we define $x \simeq y$ if $x$ and $y$ obey exactly the same formulas of $\mathcal{L}_0$.
- We claim that $\simeq$ is a bisimulation relation.
- Suppose that $x, y \in S$ and for some $a$ and some $\simeq$-closed set $C$, $\tau_a(x, C) \neq \tau_a(y, C)$.
- We need to show there is a formula on which $x, y$ disagree.
- Let $\delta = \tau_a(x, \cdot)$ and $\gamma = \tau_a(y, \cdot)$.

# Bisimulation proof I

- Given $(S, \Sigma, \tau_a)$ an LMP, we define $x \simeq y$ if $x$ and $y$ obey exactly the same formulas of $\mathcal{L}_0$.
- We claim that $\simeq$ is a bisimulation relation.
- Suppose that $x, y \in S$ and for some $a$ and some $\simeq$-closed set $C$, $\tau_a(x, C) \neq \tau_a(y, C)$.
- We need to show there is a formula on which $x, y$ disagree.
- Let $\delta = \tau_a(x, \cdot)$ and $\gamma = \tau_a(y, \cdot)$.
- If $\delta(S) > \gamma(S)$ then choose *rational* $q$ such that $\delta(S) > q > \gamma(S)$. Now $x \models \langle a \rangle_q \top$ and $y \not\models \langle a \rangle_q \top$.

# Bisimulation proof II

- If $\delta(S) = \gamma(S)$ then pick an $\simeq$-closed set $C \in \Sigma$ with $\delta(C) \neq \gamma(C)$.

# Bisimulation proof II

- If $\delta(S) = \gamma(S)$ then pick an $\simeq$-closed set $C \in \Sigma$ with $\delta(C) \neq \gamma(C)$.
- Define $\Pi = \{[\![\phi]\!] | \phi \in \mathcal{L}_0\}$ and $\Lambda = \{Y \in \Sigma | \delta(Y) = \gamma(Y)\}$. These are a $\pi$-system and a $\lambda$-system respectively.

# Bisimulation proof II

- If $\delta(S) = \gamma(S)$ then pick an $\simeq$-closed set $C \in \Sigma$ with $\delta(C) \neq \gamma(C)$.
- Define $\Pi = \{[\![\phi]\!] | \phi \in \mathcal{L}_0\}$ and $\Lambda = \{Y \in \Sigma | \delta(Y) = \gamma(Y)\}$. These are a $\pi$-system and a $\lambda$-system respectively.
- By unique structure theorem $C \in \sigma(\Pi)$ but, by assumption $C \notin \Lambda$ so $\Pi \not\subset \Lambda$ so there is a formula $\phi$ such that $\delta([\![\phi]\!]) \neq \gamma([\![\phi]\!])$.

# Bisimulation proof II

- If $\delta(S) = \gamma(S)$ then pick an $\simeq$-closed set $C \in \Sigma$ with $\delta(C) \neq \gamma(C)$.
- Define $\Pi = \{[\![\phi]\!] | \phi \in \mathcal{L}_0\}$ and $\Lambda = \{Y \in \Sigma | \delta(Y) = \gamma(Y)\}$. These are a $\pi$-system and a $\lambda$-system respectively.
- By unique structure theorem $C \in \sigma(\Pi)$ but, by assumption $C \notin \Lambda$ so $\Pi \not\subset \Lambda$ so there is a formula $\phi$ such that $\delta([\![\phi]\!]) \neq \gamma([\![\phi]\!])$.
- Suppose $\delta([\![\phi]\!]) > \gamma([\![\phi]\!])$ choose $q$ rational in between and we have

# Bisimulation proof II

- If $\delta(S) = \gamma(S)$ then pick an $\simeq$-closed set $C \in \Sigma$ with $\delta(C) \neq \gamma(C)$.
- Define $\Pi = \{[\![\phi]\!] | \phi \in \mathcal{L}_0\}$ and $\Lambda = \{Y \in \Sigma | \delta(Y) = \gamma(Y)\}$. These are a $\pi$-system and a $\lambda$-system respectively.
- By unique structure theorem $C \in \sigma(\Pi)$ but, by assumption $C \notin \Lambda$ so $\Pi \not\subset \Lambda$ so there is a formula $\phi$ such that $\delta([\![\phi]\!]) \neq \gamma([\![\phi]\!])$.
- Suppose $\delta([\![\phi]\!]) > \gamma([\![\phi]\!])$ choose $q$ rational in between and we have
- $x \models \langle a \rangle_q \phi$ and $y \not\models \langle a \rangle_q \phi$.

# How can we do this for simulation?

- Simulation is a preorder $\preceq$ rather than an equivalence relation.

# How can we do this for simulation?

- Simulation is a preorder $\preceq$ rather than an equivalence relation.
- Simulation game can be defined similarly: Duplicator starts by claiming $x \preceq y$.

# How can we do this for simulation?

- Simulation is a preorder $\preceq$ rather than an equivalence relation.
- Simulation game can be defined similarly: Duplicator starts by claiming $x \preceq y$.
- Spoiler chooses $C$ which he claims is $\preceq$-closed and that $\tau(x, C) > \tau(y, C)$.

# How can we do this for simulation?

- Simulation is a preorder $\preceq$ rather than an equivalence relation.
- Simulation game can be defined similarly: Duplicator starts by claiming $x \preceq y$.
- Spoiler chooses $C$ which he claims is $\preceq$-closed and that $\tau(x, C) > \tau(y, C)$.
- Duplicator chooses $x' \in C$ and $y' \notin C$ and claims that $x' \preceq y'$.

# How can we do this for simulation?

- Simulation is a preorder $\preceq$ rather than an equivalence relation.
- Simulation game can be defined similarly: Duplicator starts by claiming $x \preceq y$.
- Spoiler chooses $C$ which he claims is $\preceq$-closed and that $\tau(x, C) > \tau(y, C)$.
- Duplicator chooses $x' \in C$ and $y' \notin C$ and claims that $x' \preceq y'$.
- $x \preceq y$ iff Duplicator has a winning strategy starting from $x, y$.

# Positive theorems

- We had to come up with positive versions of the unique structure theorem and the monotone class theorem. With help from experts in descriptive theory.

# Positive theorems

- We had to come up with positive versions of the unique structure theorem and the monotone class theorem. With help from experts in descriptive theory.
- With these in place the proof of the logical characterization of simulation follows the same pattern.

# Uncountable labels

- The logical characterization theorem is *false* if you allow uncountably many labels. [Fijalkow]

# Uncountable labels

- The logical characterization theorem is *false* if you allow uncountably many labels. [Fijalkow]
- However, if you require the transition functions to be continuous instead of measurable then logical characterization is restored.

# Uncountable labels

- The logical characterization theorem is *false* if you allow uncountably many labels. [Fijalkow]
- However, if you require the transition functions to be continuous instead of measurable then logical characterization is restored.
- For simulation as well as bisimulation.

# Uncountable labels

- The logical characterization theorem is *false* if you allow uncountably many labels. [Fijalkow]
- However, if you require the transition functions to be continuous instead of measurable then logical characterization is restored.
- For simulation as well as bisimulation.
- We heavily use topological ideas in this proof.

# But...

- In the context of probability is exact equivalence reasonable?

# But...

- In the context of probability is exact equivalence reasonable?
- We say “no”. A small change in the probability distributions may result in bisimilar processes no longer being bisimilar though they may be very “close” in behaviour.

# But...

- In the context of probability is exact equivalence reasonable?
- We say “no”. A small change in the probability distributions may result in bisimilar processes no longer being bisimilar though they may be very “close” in behaviour.
- Instead one should have a (pseudo)metric for probabilistic processes.

# Pseudometrics

- Function $d : X \times X \longrightarrow \mathbb{R}^{\geq 0}$

# Pseudometrics

- Function $d : X \times X \longrightarrow \mathbb{R}^{\geq 0}$
- $\forall s, d(s, s) = 0$; one can have $x \neq y$ and $d(x, y) = 0$.

# Pseudometrics

- Function $d : X \times X \longrightarrow \mathbb{R}^{\geq 0}$
- $\forall s, d(s, s) = 0$; one can have $x \neq y$ and $d(x, y) = 0$.
- $\forall s, t, d(s, t) = d(t, s)$

# Pseudometrics

- Function $d : X \times X \longrightarrow \mathbb{R}^{\geq 0}$
- $\forall s, d(s, s) = 0$; one can have $x \neq y$ and $d(x, y) = 0$.
- $\forall s, t, d(s, t) = d(t, s)$
- $\forall s, t, u, d(s, u) \leq d(s, t) + d(t, u)$; triangle inequality.

# Pseudometrics

- Function $d : X \times X \longrightarrow \mathbb{R}^{\geq 0}$
- $\forall s, d(s,s) = 0$; one can have $x \neq y$ and $d(x,y) = 0$.
- $\forall s, t, d(s,t) = d(t,s)$
- $\forall s, t, u, d(s,u) \leq d(s,t) + d(t,u)$; triangle inequality.
- Quantitative analogue of an equivalence relation.

# Pseudometrics

- Function $d : X \times X \longrightarrow \mathbb{R}^{\geq 0}$
- $\forall s, d(s,s) = 0$; one can have $x \neq y$ and $d(x,y) = 0$.
- $\forall s, t, d(s,t) = d(t,s)$
- $\forall s, t, u, d(s,u) \leq d(s,t) + d(t,u)$; triangle inequality.
- Quantitative analogue of an equivalence relation.
- If we insist on $d(x,y) = 0$ iff $x = y$ we get a *metric*.

# Pseudometrics

- Function $d : X \times X \longrightarrow \mathbb{R}^{\geq 0}$
- $\forall s, d(s,s) = 0$; one can have $x \neq y$ and $d(x,y) = 0$.
- $\forall s,t, d(s,t) = d(t,s)$
- $\forall s,t,u, d(s,u) \leq d(s,t) + d(t,u)$; triangle inequality.
- Quantitative analogue of an equivalence relation.
- If we insist on $d(x,y) = 0$ iff $x = y$ we get a *metric*.
- A pseudometric defines an equivalence relation: $x \sim y$ if $d(x,y) = 0$.

# Pseudometrics

- Function $d : X \times X \longrightarrow \mathbb{R}^{\geq 0}$
- $\forall s, d(s, s) = 0$; one can have $x \neq y$ and $d(x, y) = 0$.
- $\forall s, t, d(s, t) = d(t, s)$
- $\forall s, t, u, d(s, u) \leq d(s, t) + d(t, u)$; triangle inequality.
- Quantitative analogue of an equivalence relation.
- If we insist on $d(x, y) = 0$ iff $x = y$ we get a *metric*.
- A pseudometric defines an equivalence relation: $x \sim y$ if $d(x, y) = 0$.
- Define $d^{\sim}$ on $X/\sim$ by $d^{\sim}([x], [y]) = d(x, y)$; well-defined by triangle. This is a proper metric.

# Bisimulation

- Let $R$ be an equivalence relation. $R$ is a bisimulation if: $s\ R\ t$ if $(\forall\ a)$:

$$(s \xrightarrow{a} P) \Rightarrow [t \xrightarrow{a} Q, P =_R Q]$$

$$(t \xrightarrow{a} Q) \Rightarrow [s \xrightarrow{a} P, P =_R Q]$$

# Bisimulation

- Let $R$ be an equivalence relation. $R$ is a bisimulation if: $s\ R\ t$ if $(\forall\ a)$:

$$(s \xrightarrow{a} P) \Rightarrow [t \xrightarrow{a} Q, P =_R Q]$$

$$(t \xrightarrow{a} Q) \Rightarrow [s \xrightarrow{a} P, P =_R Q]$$

- $=_R$ means that the measures $P, Q$ agree on unions of $R$-equivalence classes.

# Bisimulation

- Let $R$ be an equivalence relation. $R$ is a bisimulation if: $s\ R\ t$ if $(\forall\ a)$:

$$(s \xrightarrow{a} P) \Rightarrow [t \xrightarrow{a} Q, P =_R Q]$$

$$(t \xrightarrow{a} Q) \Rightarrow [s \xrightarrow{a} P, P =_R Q]$$

- $=_R$ means that the measures $P, Q$ agree on unions of $R$-equivalence classes.
- $s, t$ are bisimilar if there is a bisimulation relating them.

# Bisimulation

- Let $R$ be an equivalence relation. $R$ is a bisimulation if: $s\ R\ t$ if $(\forall\ a)$:

$$(s \xrightarrow{a} P) \Rightarrow [t \xrightarrow{a} Q, P =_R Q]$$

$$(t \xrightarrow{a} Q) \Rightarrow [s \xrightarrow{a} P, P =_R Q]$$

- $=_R$ means that the measures $P, Q$ agree on unions of $R$-equivalence classes.
- $s, t$ are bisimilar if there is a bisimulation relating them.
- There is a maximum bisimulation relation.

# A putative definition of a metric-bisimulation

- $m$ is a metric-bisimulation if: $m(s,t) < \epsilon \Rightarrow$:

$$s \longrightarrow P \Rightarrow \ t \longrightarrow Q, \quad m(P,Q) < \epsilon$$

$$t \longrightarrow Q \Rightarrow \ s \longrightarrow P, \ \ m(P,Q) < \epsilon$$

# A putative definition of a metric-bisimulation

- $m$ is a metric-bisimulation if: $m(s,t) < \epsilon \Rightarrow$:

$$s \longrightarrow P \Rightarrow \; t \longrightarrow Q, \quad m(P,Q) < \epsilon$$

$$t \longrightarrow Q \Rightarrow \; s \longrightarrow P, \;\; m(P,Q) < \epsilon$$

- Problem: what is $m(P,Q)$? — Type mismatch!!

# A putative definition of a metric-bisimulation

- $m$ is a metric-bisimulation if: $m(s,t) < \epsilon \Rightarrow$:

$$s \longrightarrow P \Rightarrow \; t \longrightarrow Q, \quad m(P,Q) < \epsilon$$

$$t \longrightarrow Q \Rightarrow \; s \longrightarrow P, \;\; m(P,Q) < \epsilon$$

- Problem: what is $m(P,Q)$? — Type mismatch!!
- Need a way to lift distances from states to a distances on distributions of states.

# A detour: Kantorovich metric

- Metrics on probability measures on metric spaces.

# A detour: Kantorovich metric

- Metrics on probability measures on metric spaces.
- $\mathcal{M}$: $1$-bounded pseudometrics on states.

# A detour: Kantorovich metric

- Metrics on probability measures on metric spaces.
- $\mathcal{M}$: $1$-bounded pseudometrics on states.
- $$d(\mu, \nu) = \sup_f |\int f d\mu - \int f d\nu|, f \text{ 1-Lipschitz}$$

# A detour: Kantorovich metric

- Metrics on probability measures on metric spaces.
- $\mathcal{M}$: $1$-bounded pseudometrics on states.
- $$d(\mu, \nu) = \sup_f |\int f d\mu - \int f d\nu|, f \text{ 1-Lipschitz}$$
- Arises in the solution of an LP problem: *transshipment*.

# An LP version for Finite-State Spaces

When state space is finite: Let $P, Q$ be probability distributions. Then:

$$m(P,Q) = \max \sum_i (P(s_i) - Q(s_i))a_i$$

subject to:

$$\forall i. 0 \leq a_i \leq 1$$
$$\forall i,j.\ a_i - a_j \leq m(s_i, s_j).$$

# The dual form

- Dual form from Worrell and van Breugel:

# The dual form

- Dual form from Worrell and van Breugel:
- $$\min \sum_{i,j} l_{ij} m(s_i, s_j) + \sum_i x_i + \sum_j y_j$$

  subject to:

  $$\begin{array}{l} \forall i.\ \sum_j l_{ij} + x_i = P(s_i) \\ \forall j.\ \sum_i l_{ij} + y_j = Q(s_j) \\ \forall i,j.\ l_{ij}, x_i, y_j \geq 0. \end{array}$$

# The dual form

- Dual form from Worrell and van Breugel:
- $$\min \sum_{i,j} l_{ij} m(s_i, s_j) + \sum_i x_i + \sum_j y_j$$

  subject to:

  $$\begin{array}{l} \forall i.\ \sum_j l_{ij} + x_i = P(s_i) \\ \forall j.\ \sum_i l_{ij} + y_j = Q(s_j) \\ \forall i,j.\ l_{ij}, x_i, y_j \geq 0. \end{array}$$

- We prove many equations by using the primal form to show one direction and the dual to show the other.

# Example 1

- $m(P, P) = 0$.

# Example 1

- $m(P, P) = 0$.
- In dual, match each state with itself, $l_{ij} = \delta_{ij} P(s_i), x_i = y_j = 0$. So:

$$\sum_{i,j} l_{ij} m(s_i, s_j) + \sum_i x_i + \sum_j y_j$$

becomes $0$.

# Example 1

- $m(P, P) = 0$.
- In dual, match each state with itself, $l_{ij} = \delta_{ij} P(s_i), x_i = y_j = 0$. So:

$$\sum_{i,j} l_{ij} m(s_i, s_j) + \sum_i x_i + \sum_j y_j$$

  becomes $0$.
- This clearly cannot be lowered further so this is the min.

# Example 2

- Let $m(s,t) = r < 1$. Let $\delta_s$( resp. $\delta_t$) be the probability measure concentrated at $s$(resp. $t$). Then,

$$m(\delta_s, \delta_t) = r$$

# Example 2

- Let $m(s,t) = r < 1$. Let $\delta_s$( resp. $\delta_t$) be the probability measure concentrated at $s$(resp. $t$). Then,

$$m(\delta_s, \delta_t) = r$$

- Upper bound from dual: Choose $l_{st} = 1$ all other $l_{ij} = 0$. Then

$$\sum_{ij} l_{ij} m(s_i, s_j) = m(s,t) = r.$$

# Example 2

- Let $m(s,t) = r < 1$. Let $\delta_s$( resp. $\delta_t$) be the probability measure concentrated at $s$(resp. $t$). Then,

$$m(\delta_s, \delta_t) = r$$

- Upper bound from dual: Choose $l_{st} = 1$ all other $l_{ij} = 0$. Then

$$\sum_{ij} l_{ij} m(s_i, s_j) = m(s,t) = r.$$

- Lower bound from primal: Choose $a_s = 0, a_t = r$, all others to match the constraints. Then

$$\sum_i (\delta_t(s_i) - \delta_s(s_i)) a_i = r.$$

# The Importance of Example 2

We can *isometrically* embed the original space in the metric space of distributions.

# Return from detour

## Summary

Given a metric on states in a metric space, can lift to a metric on probability distributions on states.

# Metric “bisimulation”

- $m$ is a metric-bisimulation if: $m(s,t) < \epsilon \Rightarrow$:

$$s \longrightarrow P \Rightarrow \; t \longrightarrow Q, \quad m(P,Q) < \epsilon$$

$$t \longrightarrow Q \Rightarrow \; s \longrightarrow P, \;\; m(P,Q) < \epsilon$$

# Metric “bisimulation”

- $m$ is a metric-bisimulation if: $m(s,t) < \epsilon \Rightarrow$:

$$s \longrightarrow P \Rightarrow \; t \longrightarrow Q, \quad m(P,Q) < \epsilon$$

$$t \longrightarrow Q \Rightarrow \; s \longrightarrow P, \;\; m(P,Q) < \epsilon$$

- The required canonical metric on processes is the least such: ie. the distances are the least possible.

# Metric "bisimulation"

- $m$ is a metric-bisimulation if: $m(s,t) < \epsilon \Rightarrow$:

$$s \longrightarrow P \Rightarrow \; t \longrightarrow Q, \quad m(P,Q) < \epsilon$$

$$t \longrightarrow Q \Rightarrow \; s \longrightarrow P, \;\; m(P,Q) < \epsilon$$

- The required canonical metric on processes is the least such: ie. the distances are the least possible.
- Thm: *Canonical least metric exists*.

# Tarski's theorem

If $L$ is a complete lattice and $F : L \longrightarrow L$ is monotone then the set of fixed points of $F$ with the induced order is itself a complete lattice. In particular there is a least fixed point and a greatest fixed point.

# Metrics: some details

- $\mathcal{M}$: $1$-bounded pseudometrics on states with ordering

$$m_1 \preceq m_2 \text{ if } (\forall s, t)\ [m_1(s,t) \geq m_2(s,t)]$$

# Metrics: some details

- $\mathcal{M}$: $1$-bounded pseudometrics on states with ordering

$$m_1 \preceq m_2 \text{ if } (\forall s, t)\ [m_1(s,t) \geq m_2(s,t)]$$

- $(\mathcal{M}, \preceq)$ is a complete lattice.

# Metrics: some details

- $\mathcal{M}$: $1$-bounded pseudometrics on states with ordering

$$m_1 \preceq m_2 \text{ if } (\forall s, t)\ [m_1(s,t) \geq m_2(s,t)]$$

- $(\mathcal{M}, \preceq)$ is a complete lattice.
-

$$\begin{aligned}
\bot(s,t) &= \begin{cases} 0 \text{ if } s = t \\ 1 \text{ otherwise} \end{cases} \\
\top(s,t) &= 0, (\forall s, t) \\
(\sqcap\{m_i\}(s,t) &= \sup_i m_i(s,t)
\end{aligned}$$

# Greatest fixed-point definition

- Let $m \in \mathcal{M}$. $F(m)(s,t) < \epsilon$ if:

$$s \longrightarrow P \Rightarrow \ t \longrightarrow Q, \quad m(P,Q) < \epsilon$$

$$t \longrightarrow Q \Rightarrow \ s \longrightarrow P, \ \ m(P,Q) < \epsilon$$

# Greatest fixed-point definition

- Let $m \in \mathcal{M}$. $F(m)(s,t) < \epsilon$ if:

$$s \longrightarrow P \Rightarrow \; t \longrightarrow Q, \quad m(P,Q) < \epsilon$$

$$t \longrightarrow Q \Rightarrow \; s \longrightarrow P, \quad m(P,Q) < \epsilon$$

- $F(m)(s,t)$ can be given by an explicit expression.

# Greatest fixed-point definition

- Let $m \in \mathcal{M}$. $F(m)(s,t) < \epsilon$ if:

$$s \longrightarrow P \Rightarrow \; t \longrightarrow Q, \quad m(P,Q) < \epsilon$$

$$t \longrightarrow Q \Rightarrow \; s \longrightarrow P, \quad m(P,Q) < \epsilon$$

- $F(m)(s,t)$ can be given by an explicit expression.
- $F$ is monotone on $\mathcal{M}$, and metric-bisimulation is the greatest fixed point of $F$.

# Greatest fixed-point definition

- Let $m \in \mathcal{M}$. $F(m)(s,t) < \epsilon$ if:

$$s \longrightarrow P \Rightarrow \ t \longrightarrow Q, \quad m(P,Q) < \epsilon$$

$$t \longrightarrow Q \Rightarrow \ s \longrightarrow P, \quad m(P,Q) < \epsilon$$

- $F(m)(s,t)$ can be given by an explicit expression.
- $F$ is monotone on $\mathcal{M}$, and metric-bisimulation is the greatest fixed point of $F$.
- The closure ordinal of F is $\omega$.

# Kantorovich-Rubinstein duality

## Definition

Given two probability measures $P_1, P_2$ on $(X, \Sigma)$, a *coupling* is a measure $Q$ on the product space $X \times X$ such that the marginals are $P_1, P_2$. Write $\mathcal{C}(P_1, P_2)$ for the set of couplings between $P_1, P_2$.

# Kantorovich-Rubinstein duality

## Definition

Given two probability measures $P_1, P_2$ on $(X, \Sigma)$, a *coupling* is a measure $Q$ on the product space $X \times X$ such that the marginals are $P_1, P_2$. Write $\mathcal{C}(P_1, P_2)$ for the set of couplings between $P_1, P_2$.

## Theorem

Let $(X, d)$ be a compact metric space. Let $P_1, P_2$ be Borel probability measures on $X$

$$\sup_{f:X \longrightarrow [0,1] \text{ nonexpansive}} \left\{ \int_X f \mathrm{d}P_1 - \int_X f \mathrm{d}P_2 \right\} = \inf_{Q \in \mathcal{C}(P_1,P_2)} \left\{ \int_{X \times X} d \ \ \mathrm{d}Q \right\}$$